

GenePattern Notebook



Integrative Genomics Viewer

Jill P. Mesirov, Ph.D.

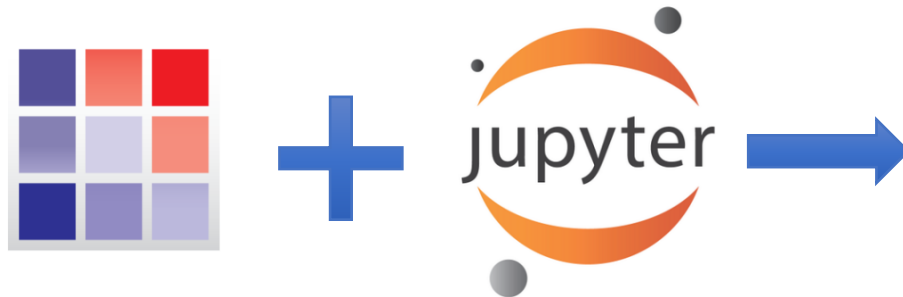
Jim Robinson, Ph.D.

ITCR PI Meeting, Santa Cruz

May 31, 2017

The GenePattern Notebook Environment

for collaborative and reproducible *in silico* research



- Shareable, executable notebook documents
- Support *in silico* research from inception through publication
- Seamless integration of GenePattern with Jupyter notebook
- No programming required
- www.genepattern-notebook.org

GenePattern Notebook Update

- GenePattern Notebook online repository released on Amazon cloud
 - Zero-install: create and run notebooks using a web browser
 - Disseminate notebooks for use and adaptation by other
- Deepened integration between code cells and GenePattern GUI cells
 - Python variables as inputs/outputs to GenePattern analyses
- Reich et al., The GenePattern Notebook environment (*Cell Systems in press*).

Acknowledgements

GenePattern Team



Peter Carr



Bobbie Hill



Ted Liefeld



Thorin Tabor



Michael Reich



Helga Thorvaldsdóttir

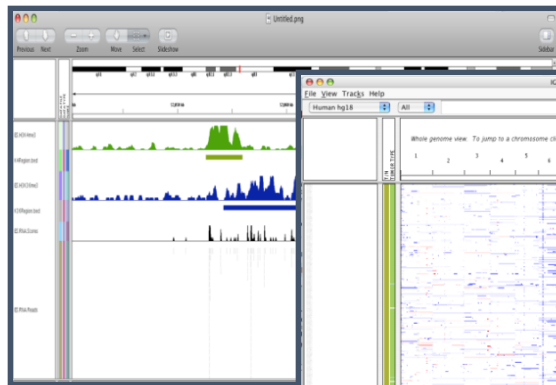


Pablo Tamayo

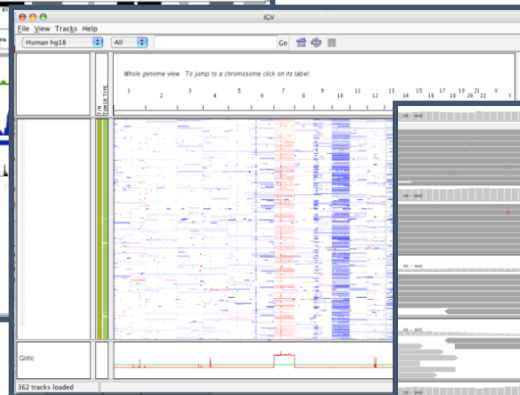
Integrative Genomics Viewer (IGV)



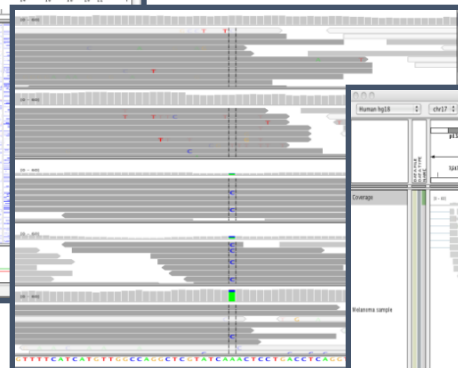
Desktop application for the visual interactive exploration of **integrated** genomic datasets



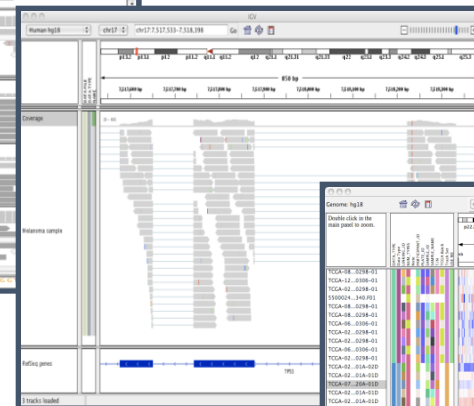
Epigenetics



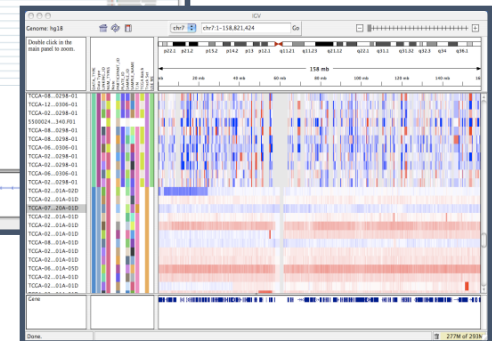
Copy Number



WG Alignments



RNA-Seq



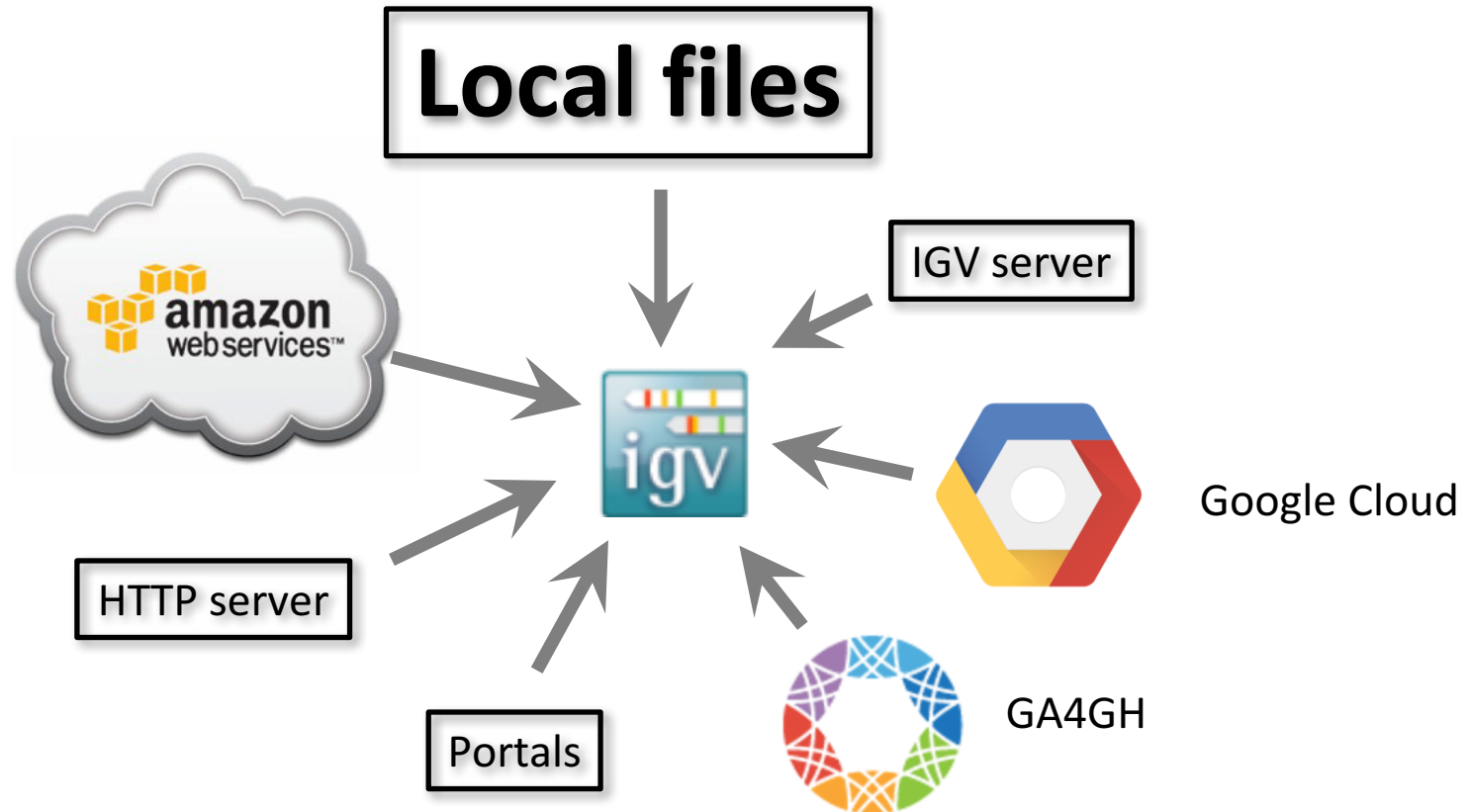
Microarrays

<https://igv.org/>

J. Robinson et al. Nature Biotech 2011

Thorvaldsdóttir et al. Briefings in Bioinformatics, 2013

IGV Data Sources

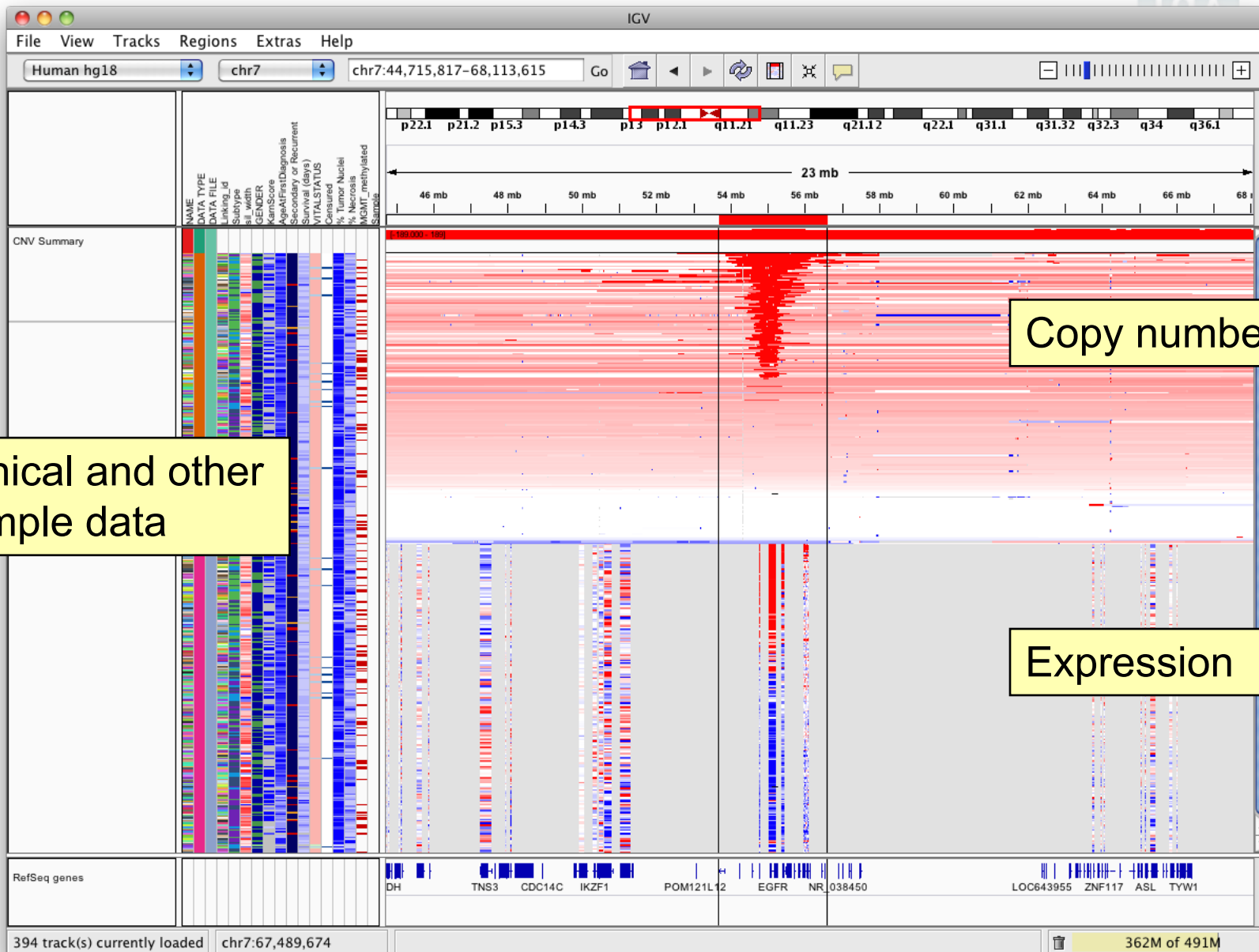


- View **local** files without uploading.
- View **remote** files without downloading the whole dataset.

Authentication

- OAuth 2
 - * *Tested with Google services and Microsoft Azure*
 - HTTP/HTTPS servers
 - Google Cloud Storage
 - GA4GH
- Signed URLs
 - Amazon Web Services

Integrated datasets

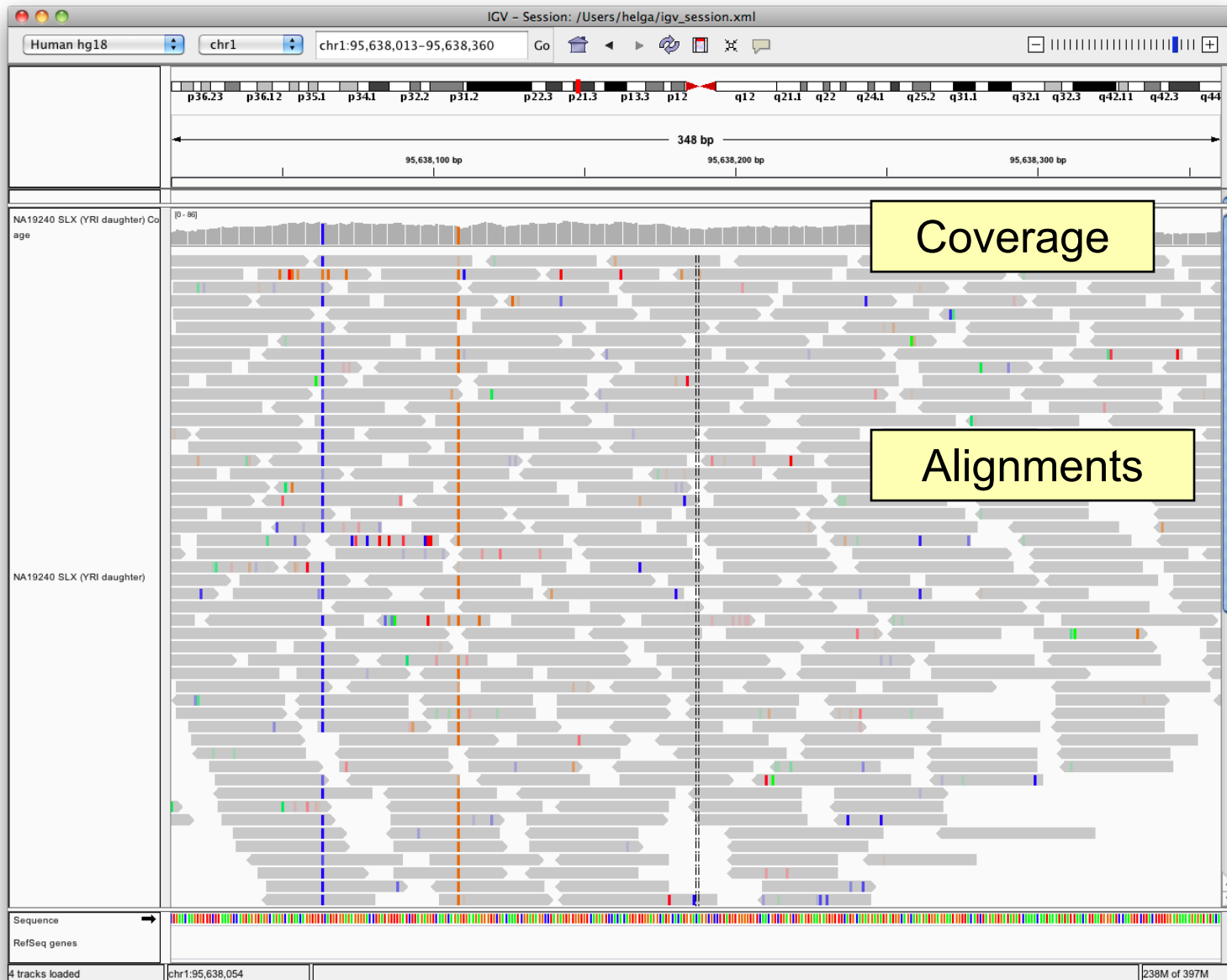


Clinical and other
sample data

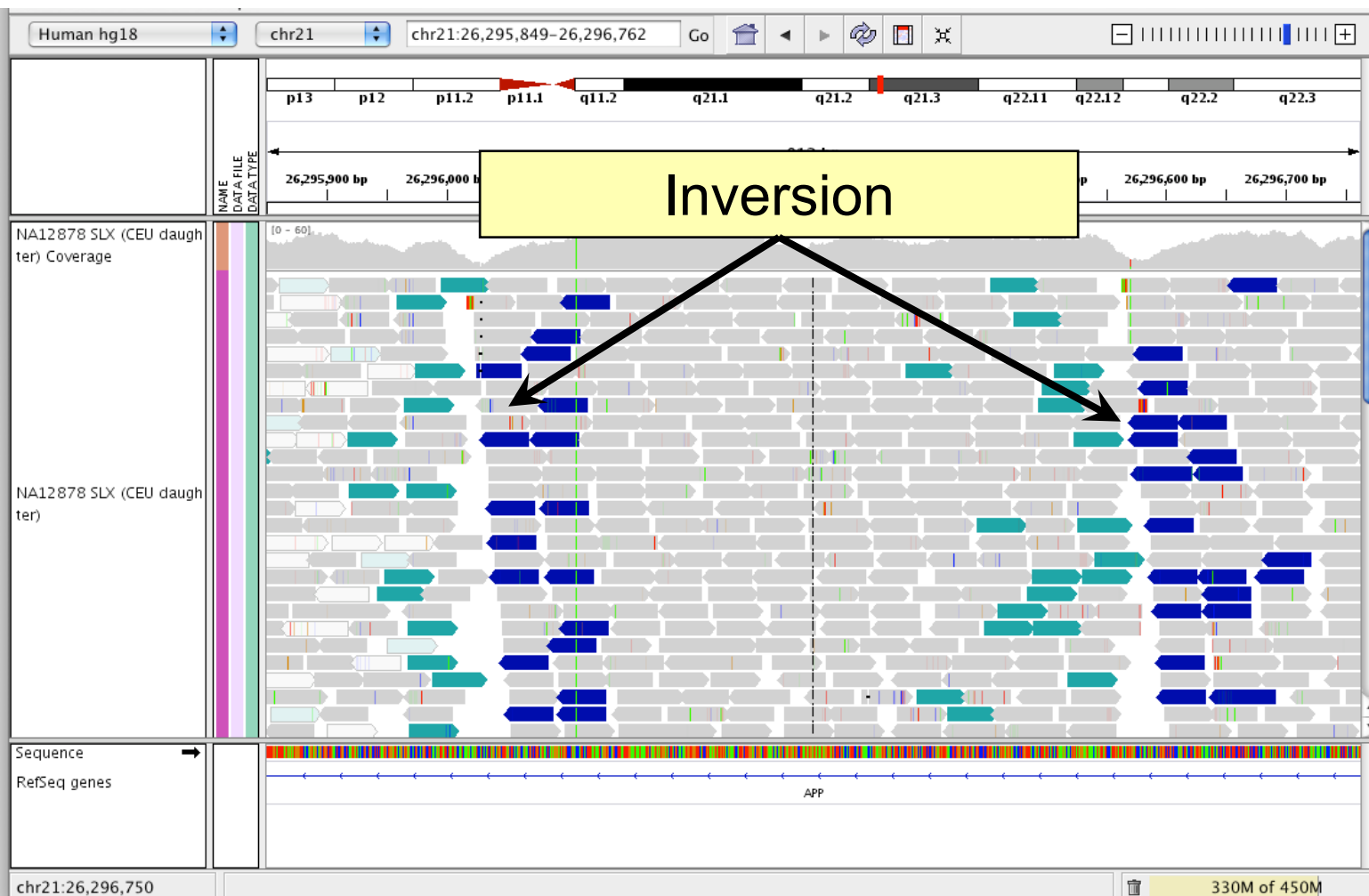
Copy number

Expression

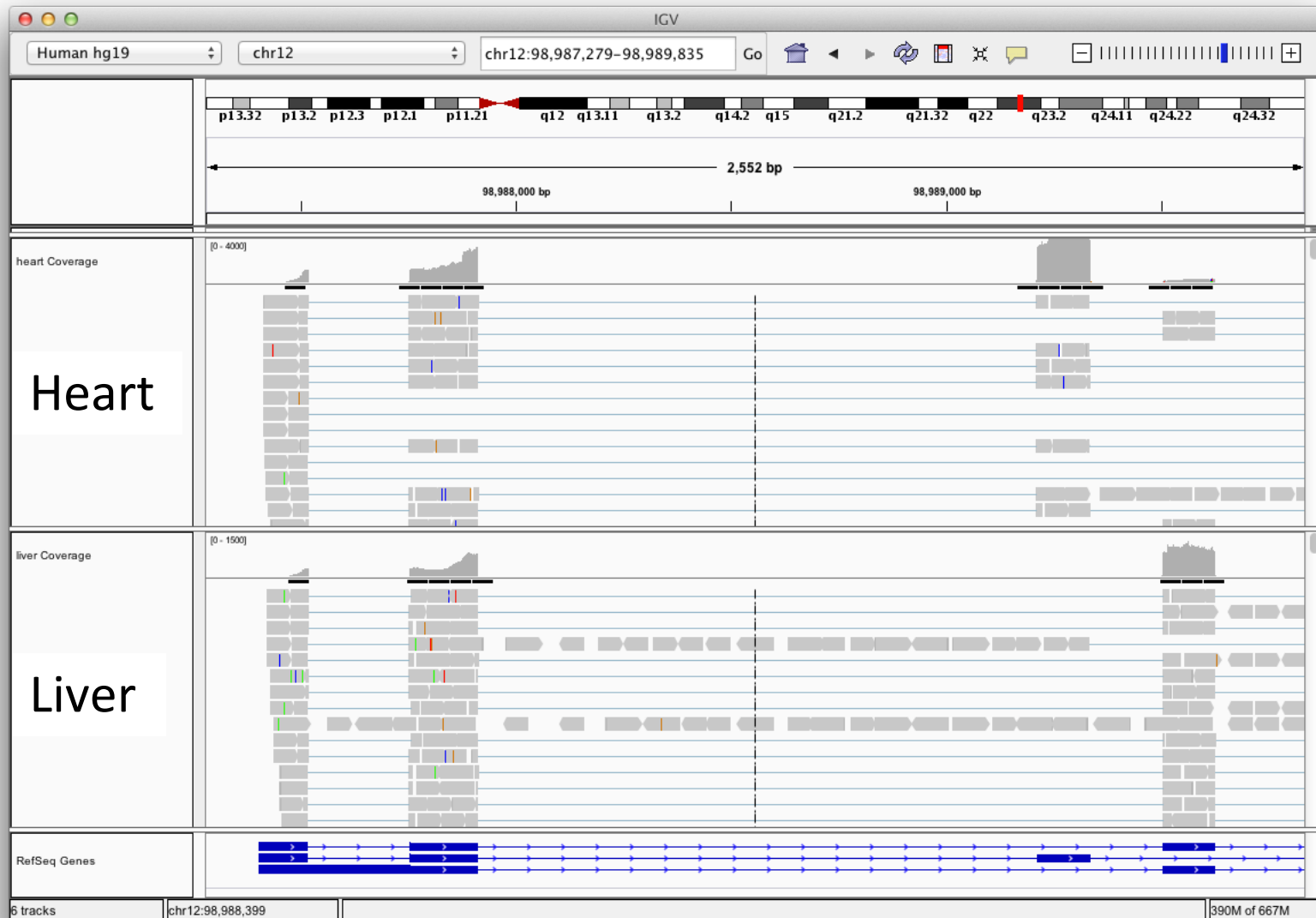
Short Read Alignments



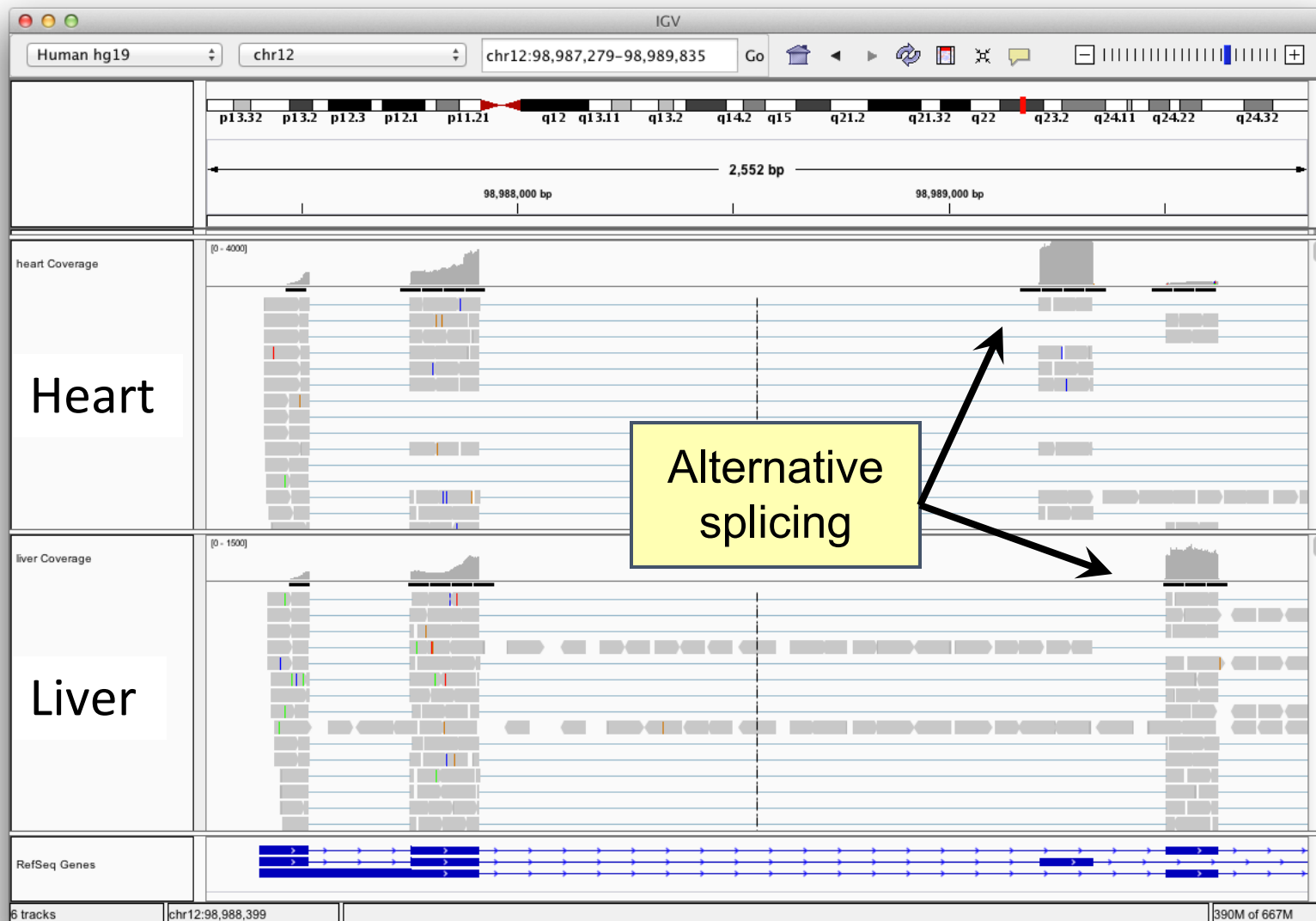
Paired-end Alignments



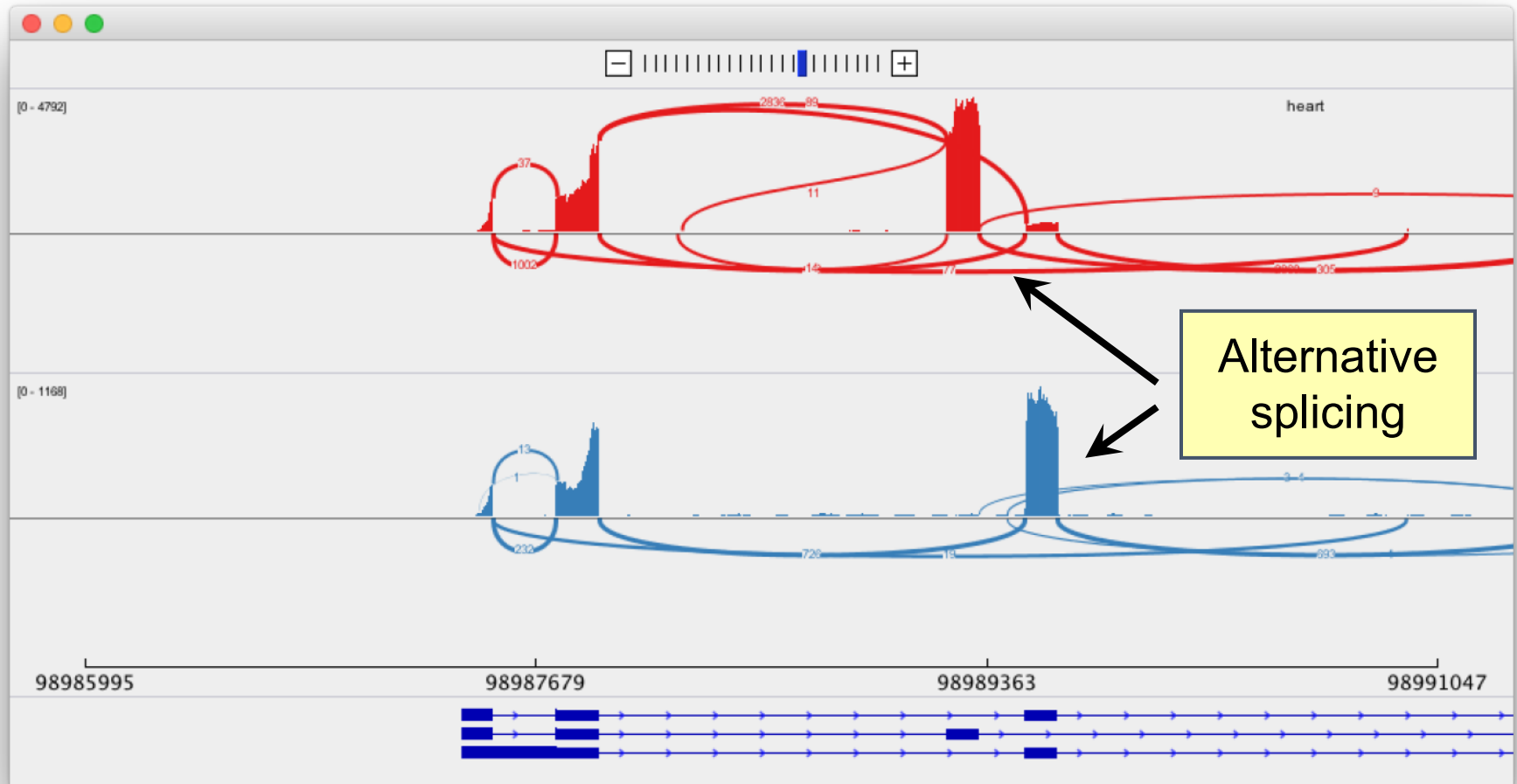
RNA-Seq



RNA-Seq



Sashimi plot

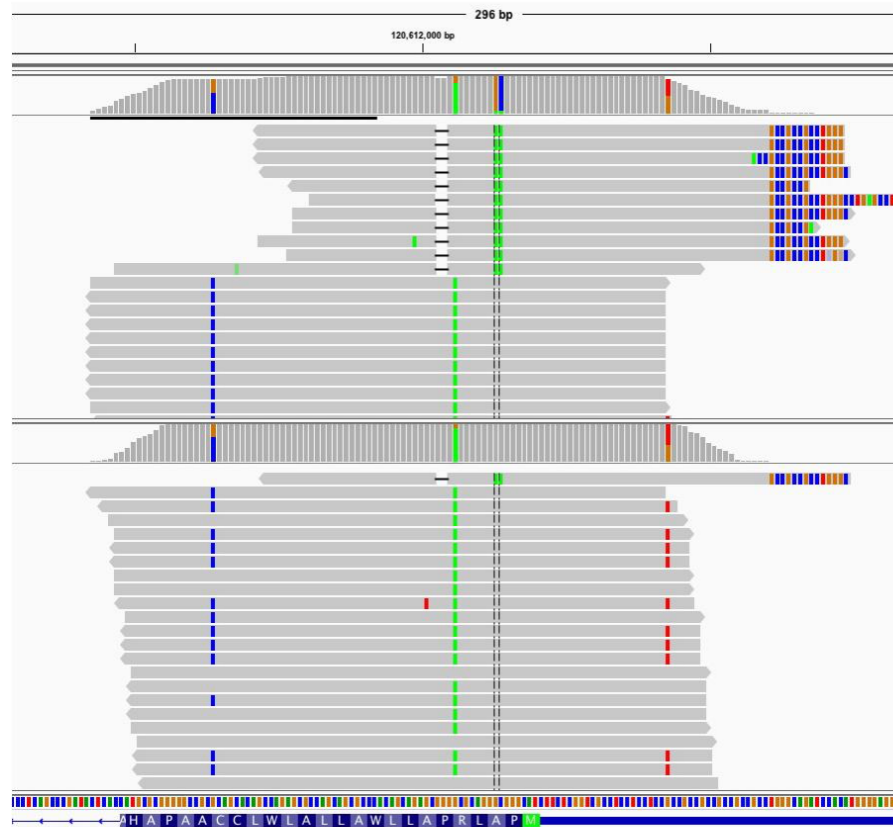


Katz et al, Quantitative visualization of alternative exon expression from RNA-seq data, Bioinformatics (2015)

Clinical example

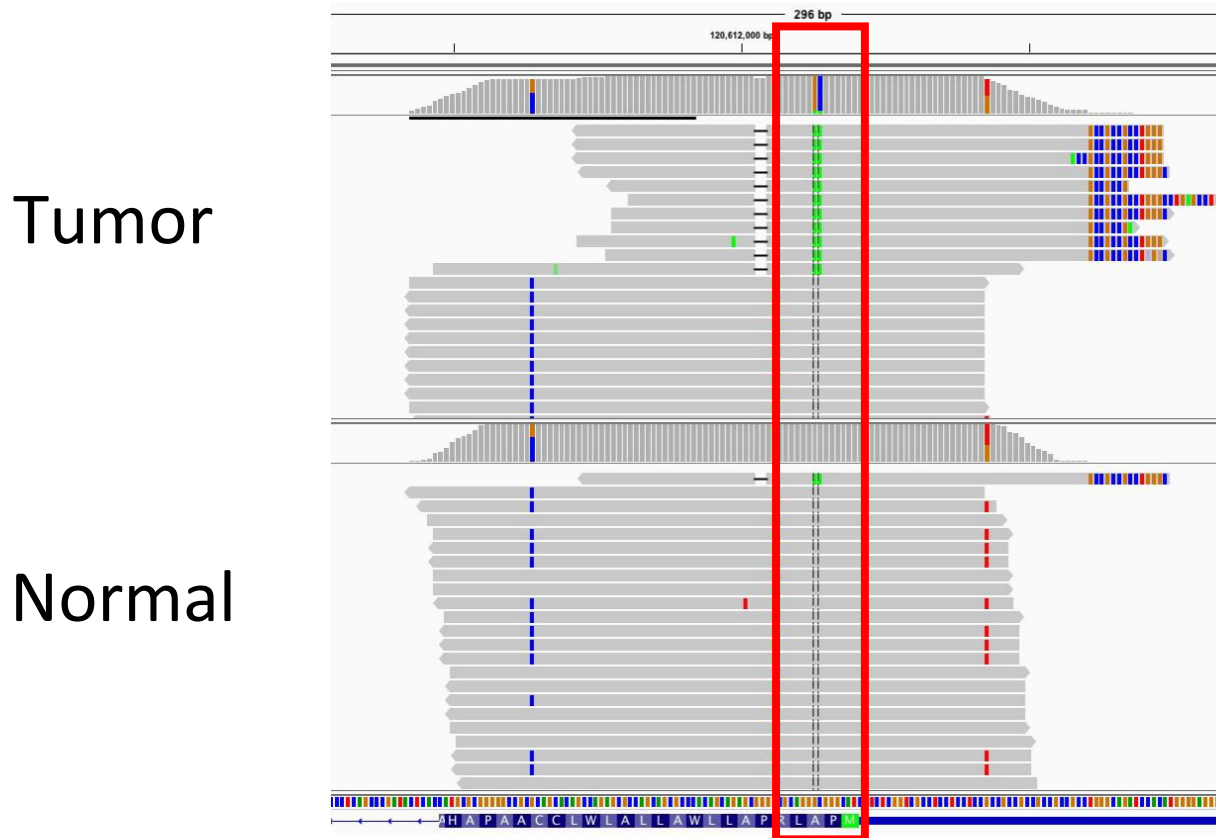
Tumor

Normal



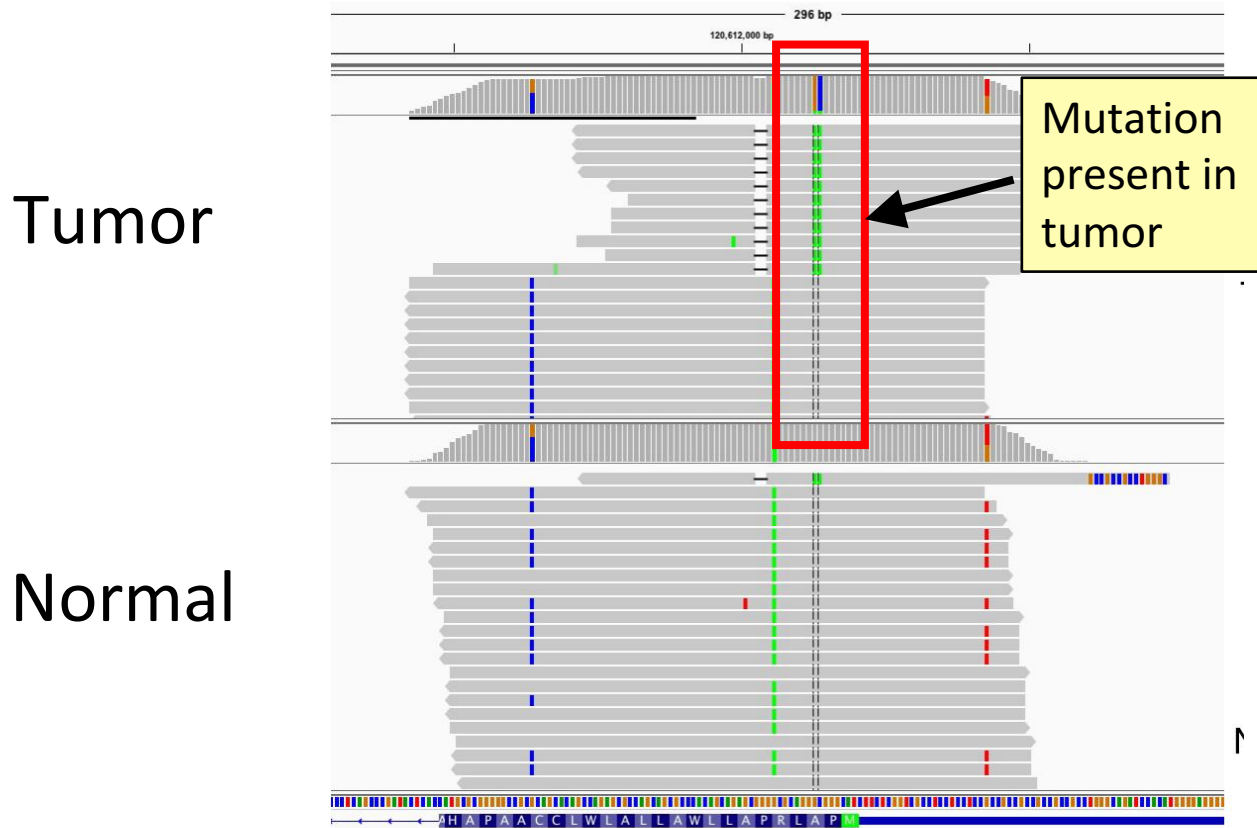
Clinical example

Pipeline call: somatic mutation



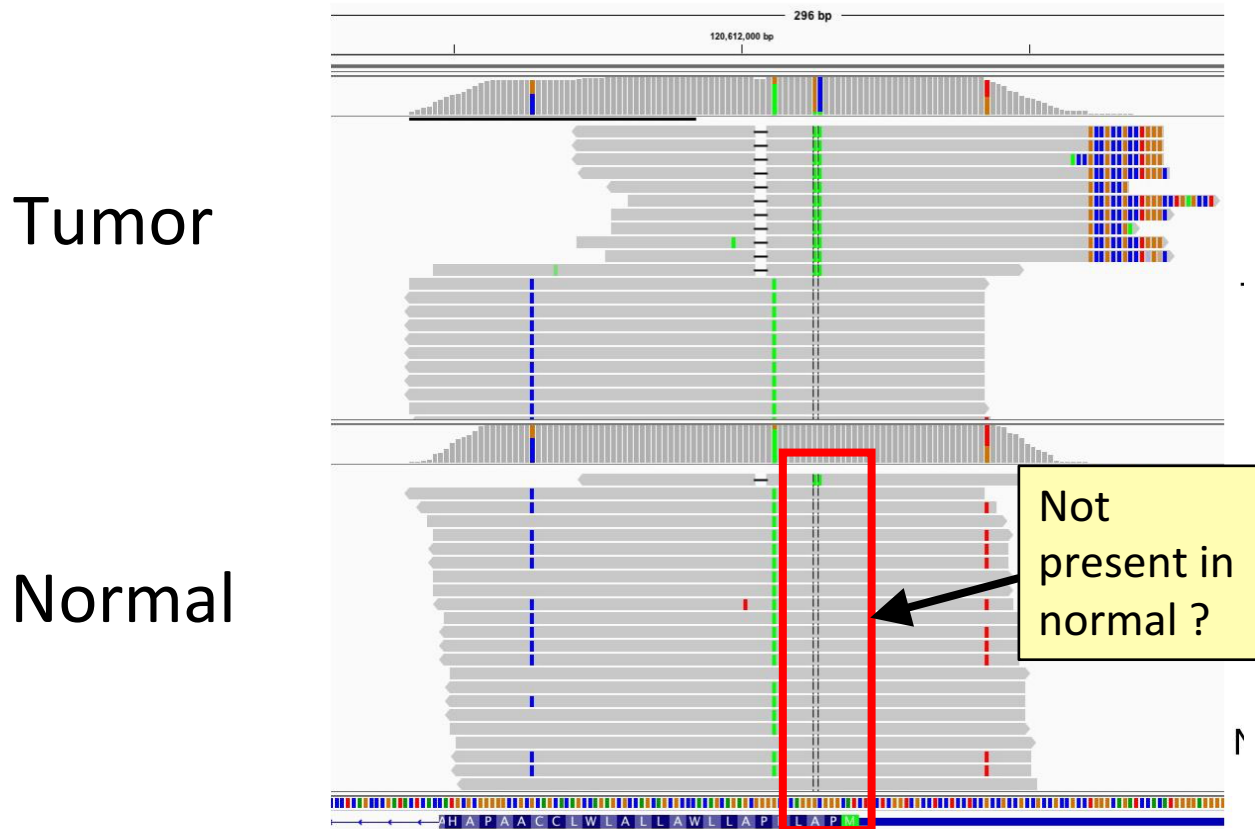
Clinical example

Pipeline call: somatic mutation



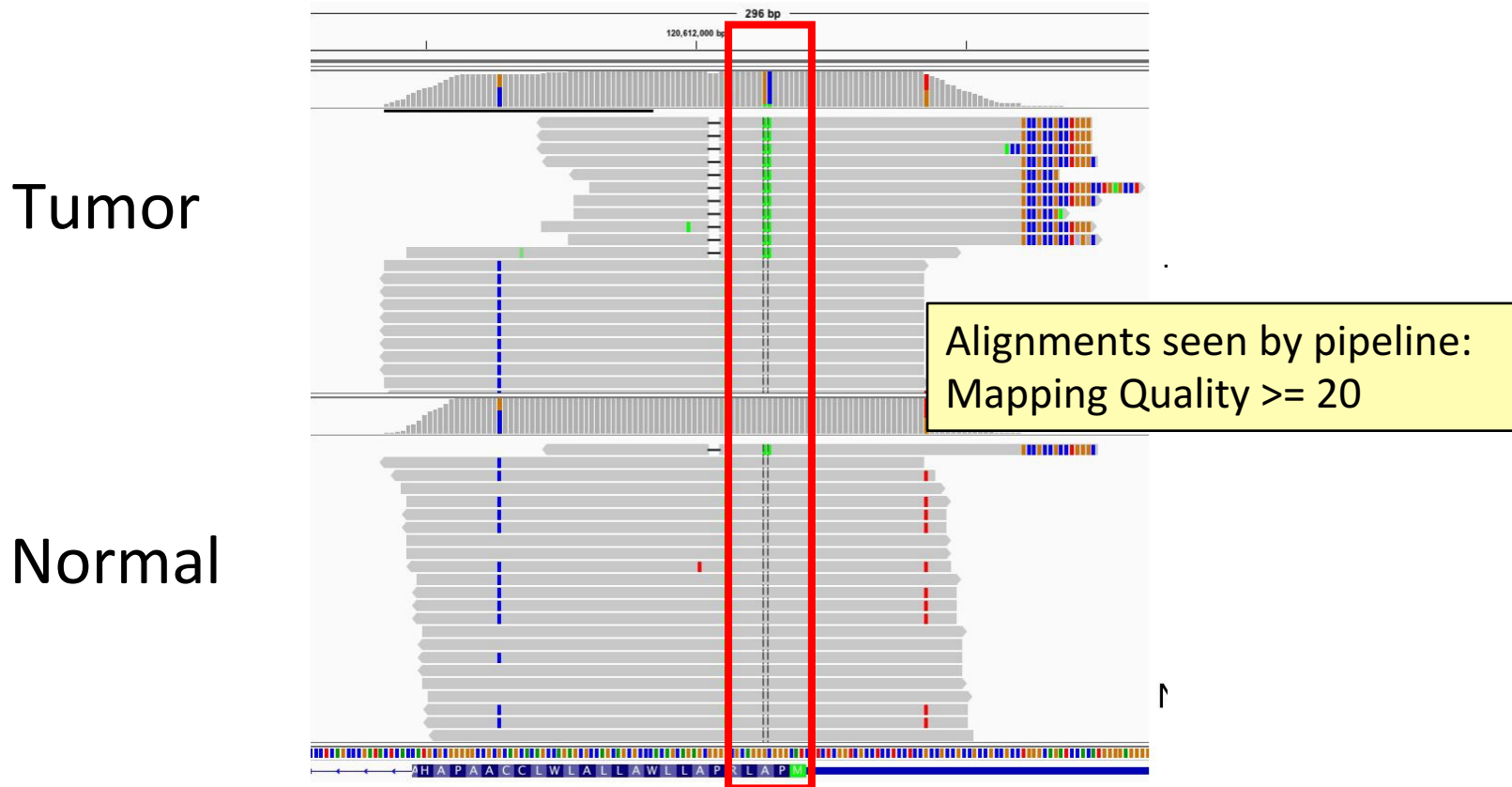
Clinical example

Pipeline call: somatic mutation



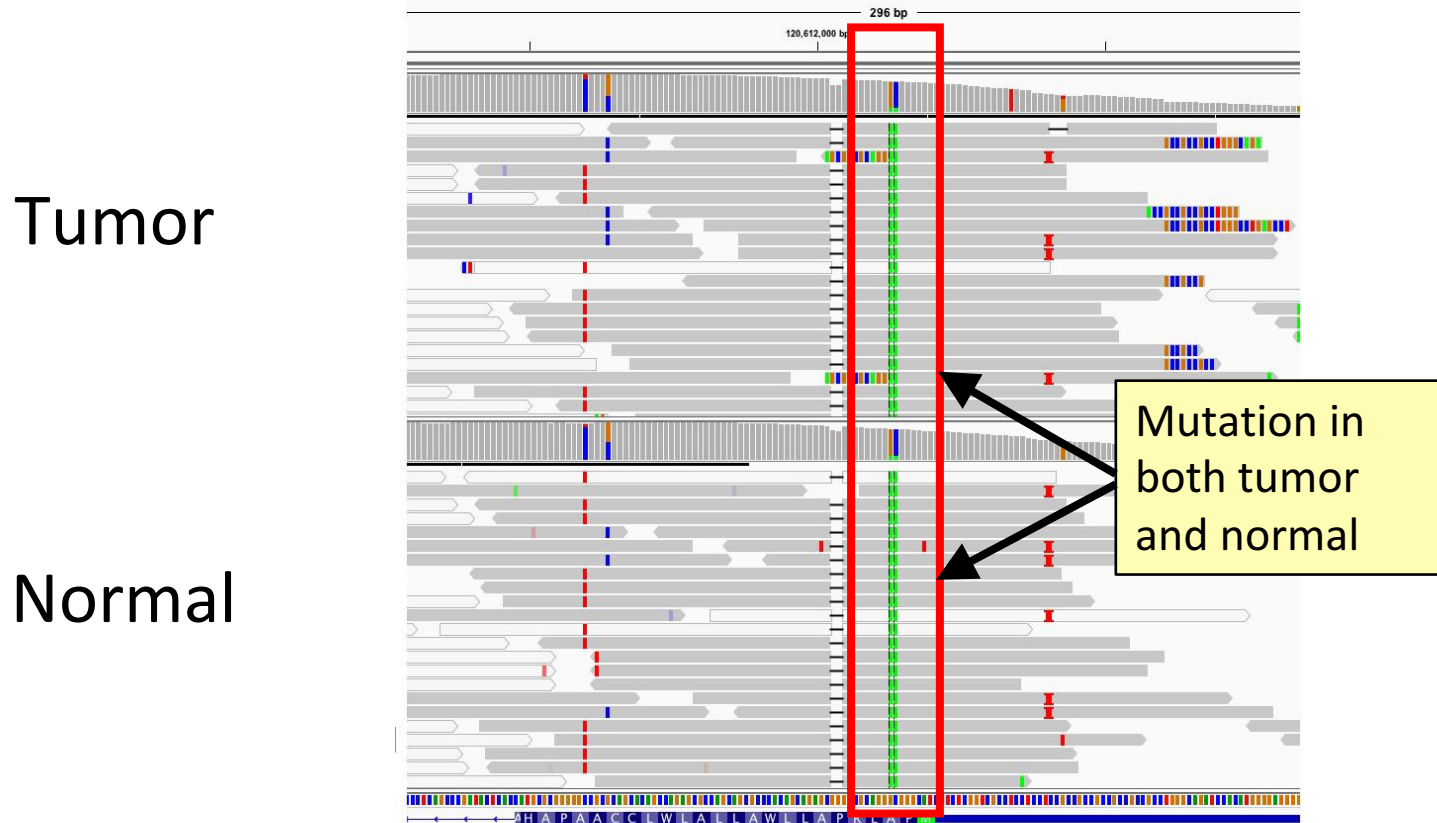
Clinical example

Pipeline call: somatic mutation



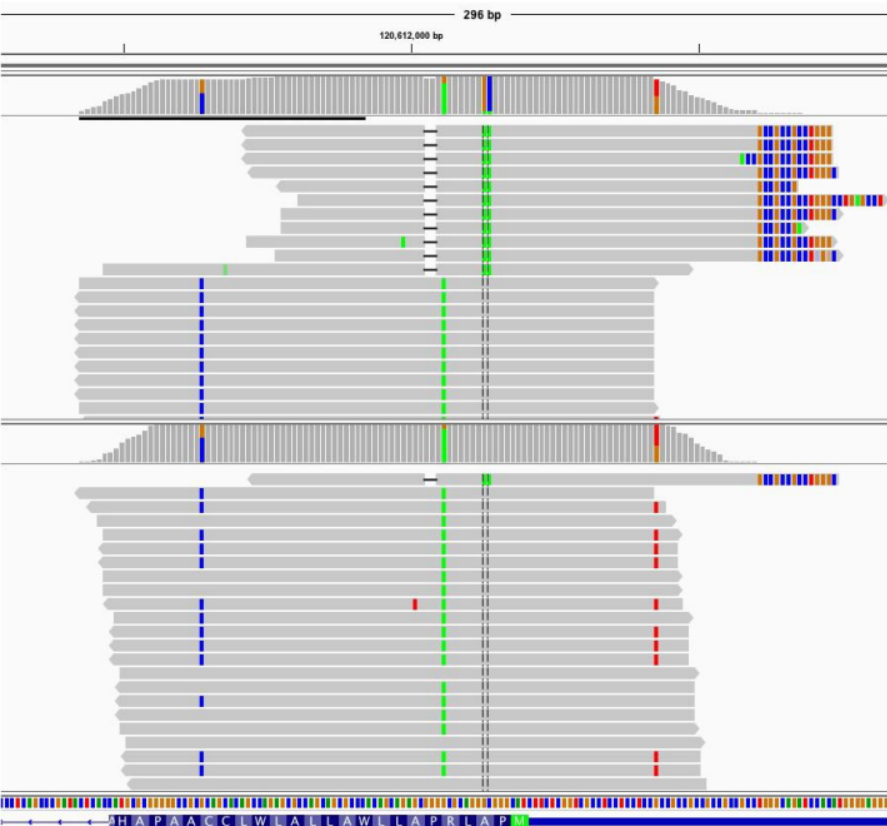
Clinical example

Pipeline call: somatic mutation



Clinical example

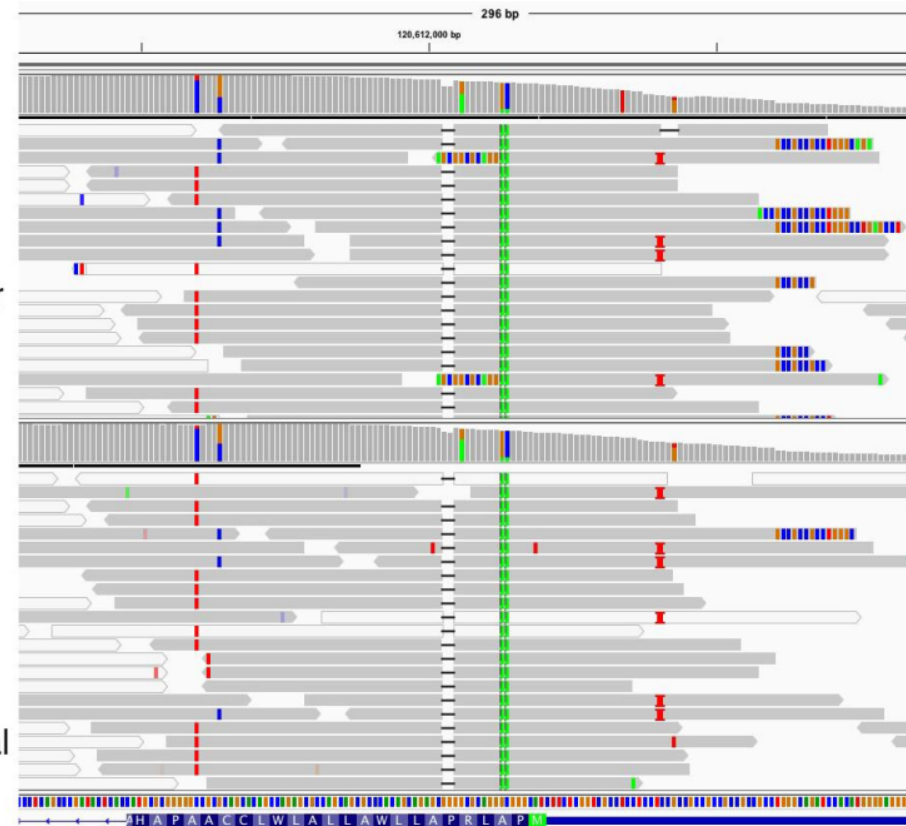
Somatic mutation cannot be confirmed.



Mapping quality threshold = 20

Tumor

Normal



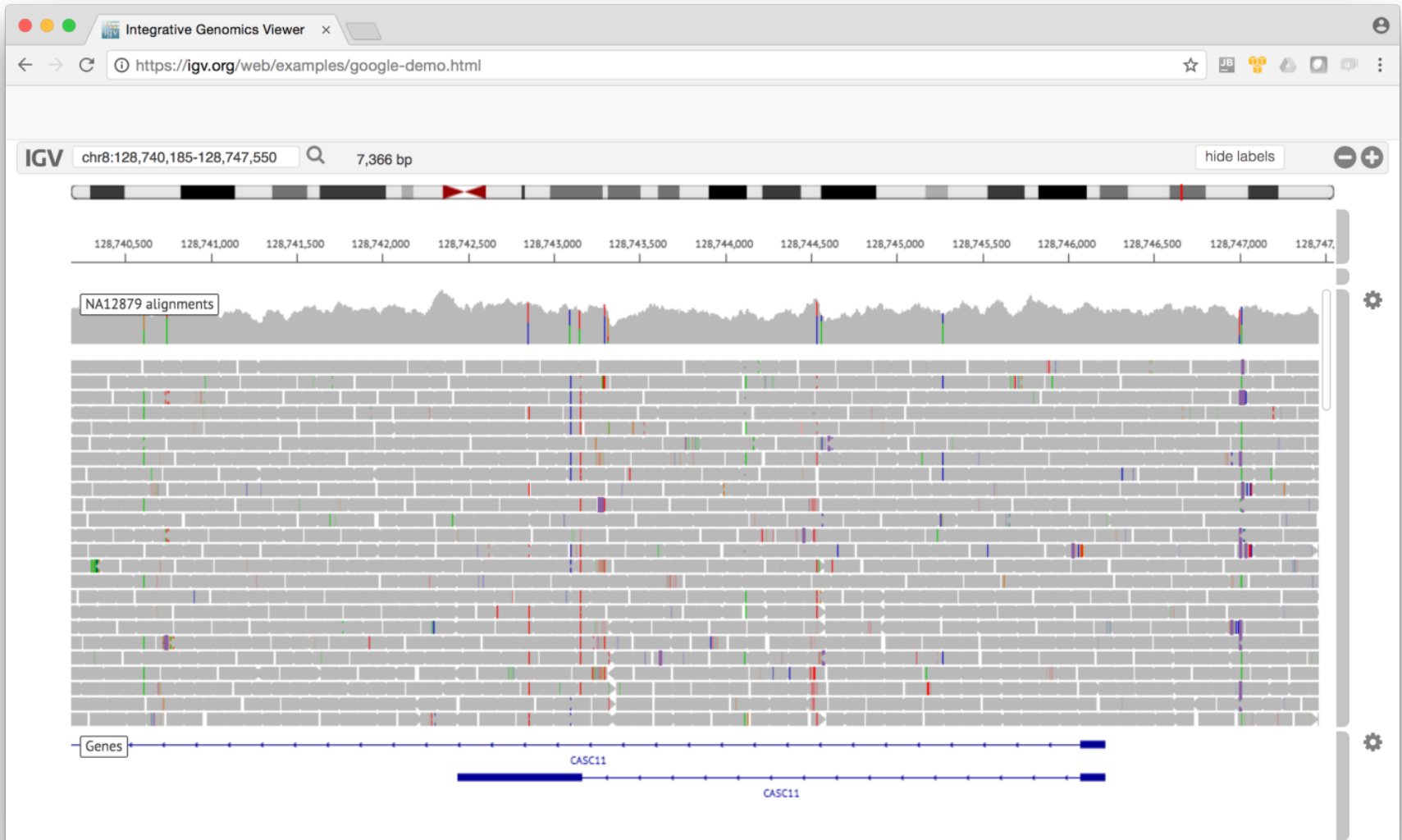
Mapping quality threshold = 0

Recent Updates



- New and enhanced features for 3rd gen sequencing (PacBio and Oxford Nanopore)
- Support for synthetic long reads (10X Genomics and Illumina Moleculo)
- Support for the CRAM alignment format
- Performance improvements
- New collaboration with Karchin Lab on MuPIT / CRAVAT integration.

- Lightweight version of IGV for embedding in web pages and portals
- Pure client component (JavaScript)
- Supports
 - All major genomic file formats
 - GA4GH API
 - Google Cloud Storage
 - Google BigTable
 - Local files
- Authentication via OAuth 2



Cloud portal integrations

- ISB Cancer Genomics Cloud
- FireCloud
- cBioPortal for Cancer Genomics
- Trinity (NCGAS Galaxy)

igv.js 1.1

- Multi-locus view
- Local file support
- New API for generic OAuth 2 authentication

IGV Usage



Since October, 2011

- 5,955,561 launches
- 490,989 unique IP addresses

Current weekly rate (average of last 4 weeks)

- 37,090 launches
- 9,726 unique IP addresses
- 2,828 new IP addresses

IGV help forum: 1,900 topics

GitHub <https://github.com/igvteam/igv>

- 111 forks
- 81 accepted pull requests
- 27 contributors, 22 external to IGV team

Citations: ~3,700

Data Credits

PacBio

<http://www.pacb.com/>

1000 Genomes

<http://www.internationalgenome.org/>

Nattestad M, Chin CS, Schatz MC. Ribbon: ***Visualizing complex genome alignments and structural variation***. bioRxiv. 2016 Jan 1:082123.

Acknowledgments



IGV Team

Helga Thorvaldsdóttir

Doug Turner

Isaac Shanley

Jim Robinson

Collaborators

Karchin Lab, John Hopkins

Regev Lab, Broad Institute

Institute for Systems Biology

cBio Team, MSKCC

Aiden Lab, Baylor

Contributors

Ben Berman

Fabian Campagne

Cassie Doll

Brian Haas

Nils Homer

Bryan LaJoie

Damon May

David Meade

Chantal Roth

Aaron Wenger